## **Claims**

What is claimed is:

5 A method for use in finding near-neighbors in a set of objects comprising the steps of:

identifying subspace pattern similarities that the objects in the set exhibit in multi-dimensional spaces; and

defining subspace correlations between two or more of the objects in the set based on the identified subspace pattern similarities for use in identifying near-neighbor objects.

- 2. The method of claim 1, wherein the identifying step further comprises the step of creating a pattern distance index.
- 3. The method of claim 1, wherein the multi-dimensional spaces comprise arbitrary spaces.
- 4. The method of claim 2, wherein the creating step further comprises the step of determining a subspace dimensionality of one or more patterns in the pattern distance index.
  - 5. The method of claim 4, wherein the subspace dimensionality is an indicator of a degree of similarity between the objects.
    - 6. The method of claim 1, wherein data relating to the objects is static.

25

10

- 7. The method of claim 1, wherein data relating to the objects comprises dynamic data insertions.
- 8. The method of claim 1, wherein data relating to the objects comprises 5 gene expression data.
  - 9. The method of claim 1, wherein data relating to the objects comprises synthetic data.
- 10. The method of claim 1, wherein identifying the subspace pattern similarities comprises a comparison of any subset of dimensions in the multi-dimensional spaces.
- 11. The method of claim 1, wherein identifying the subspace pattern similarities comprises an ordering of dimensions in the multi-dimensional spaces.
  - 12. The method of claim 1, wherein each object is represented by a sequence of pairs, each pair indicating a dimension and an object value in that dimension.
- 20 13. The method of claim 12, wherein a first pair in the sequence of pairs comprises a base of comparison for one or more remaining pairs in the sequence of pairs.
  - 14. The method of claim 12, wherein the sequence of pairs is represented sequentially in a tree structure comprising one or more edges and one or more nodes.
  - 15. The method of claim 2, wherein creating the pattern distance index comprises use of pattern-distance links.

- 16. The method of claim 1, wherein the process is optimized by maintaining a set of embedded ranges.
- 5 17. The method of claim 1, wherein the subspace correlations comprise a distance between two or more of the objects in the set.
  - 18. A method of performing a near-neighbor search of one or more query objects against a set of objects comprising the steps of:
- creating a pattern distance index to identify subspace pattern similarities that the objects in the set exhibit in multi-dimensional spaces;

defining subspace correlations between two or more of the objects in the set based on the identified subspace pattern similarities; and

using the subspace correlations to identify near-neighbor objects among
the query objects and the objects in the set.

19. An apparatus for use in finding near-neighbors in a set of objects, the apparatus comprising:

a memory; and

at least one processor, coupled to the memory, operative to:

identify subspace pattern similarities that the objects in the set exhibit in multi-dimensional spaces; and

define subspace correlations between two or more of the objects in the set based on the identified subspace pattern similarities for use in identifying near-neighbor objects.

20. An article of manufacture for finding near-neighbors in a set of objects, comprising a machine readable medium containing one or more programs which when executed implement the steps of:

identifying subspace pattern similarities that the objects in the set exhibit in multi-dimensional spaces; and

defining subspace correlations between two or more of the objects in the set based on the identified subspace pattern similarities for use in identifying near-neighbor objects.